

SEQUENCE LISTING

B1
 <110> Ago, Hideo
 Miyano, Masashi
 Adachi, Tsuyoshi

<120> HCV Polymerase Suitable for Crystal
 Structure Analysis and Method for Using the Enzyme

<130> SHIM007

<140> 09/608,713

<141> 2000-06-30

<150> 11-188630

<151> 1999-07-02

<150> 11-192488

<151> 1999-07-07

<160> 12

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 591

<212> PRT

<213> Hepatitis C Virus

<400> 1

Ser	Met	Ser	Tyr	Thr	Trp	Thr	Gly	Ala	Leu	Ile	Thr	Pro	Cys	Ala	Ala	1	5	10	15
Glu	Glu	Ser	Lys	Leu	Pro	Ile	Asn	Ala	Leu	Ser	Asn	Ser	Leu	Leu	Arg	20	25	30	35
His	His	Asn	Met	Val	Tyr	Ala	Thr	Ser	Arg	Ser	Ala	Gly	Leu	Arg		35	40	45	50
Gln	Lys	Lys	Val	Thr	Phe	Asp	Arg	Leu	Gln	Val	Leu	Asp	Asp	His	Tyr	55	60	65	70
Arg	Asp	Val	Leu	Lys	Glu	Met	Lys	Ala	Lys	Ala	Ser	Thr	Val	Lys	Ala	75	80	85	90
Lys	Leu	Leu	Ser	Val	Glu	Glu	Ala	Cys	Lys	Leu	Thr	Pro	Pro	His	Ser	95	100	105	110
Ala	Lys	Ser	Lys	Phe	Gly	Tyr	Gly	Ala	Lys	Asp	Val	Arg	Asn	Leu	Ser	115	120	125	130
Ser	Lys	Ala	Val	Asn	His	Ile	His	Ser	Val	Trp	Lys	Asp	Leu	Leu	Glu	135	140	145	150
Asp	Thr	Val	Thr	Pro	Ile	Asp	Thr	Thr	Ile	Met	Ala	Lys	Asn	Glu	Val	155	160	165	170
Phe	Cys	Val	Gln	Pro	Glu	Lys	Gly	Gly	Arg	Lys	Pro	Ala	Arg	Leu	Ile	175			
Val	Phe	Pro	Asp	Leu	Gly	Val	Arg	Val	Cys	Glu	Lys	Met	Ala	Leu	Tyr				

Asp Val Val Ser Thr Leu Pro Gln Val Val Met Gly Ser Ser Tyr Gly
180 185 190
Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Thr Trp
195 200 205
Lys Ser Lys Lys Asn Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe
210 215 220
Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile Tyr
225 230 235 240
Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Lys Ser Leu
245 250 255
Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly Gln
260 265 270
Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser
275 280 285
Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys Arg
290 295 300
Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Asn Gly Asp Asp Leu
305 310 315 320
Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser Leu
325 330 335
Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp
340 345 350
Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser
355 360 365
Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr Leu
370 375 380
Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Trp Glu Thr Ala
385 390 395 400
Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr Ala
405 410 415
Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser Ile
420 425 430
Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Gln Ile Tyr
435 440 445
Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile Glu
450 455 460
Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro Gly
465 470 475 480
Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro Pro
485 490 495
Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu
500 505 510
Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn Trp
515 520 525
Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser Gln
530 535 540
Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp Ile
545 550 555 560
Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Met Leu Cys Leu
565 570 575
Leu Leu Leu Ser Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg
580 585 590

<210> 2
<211> 2889
<212> DNA
<213> Artificial Sequence

<220>
<223> CDS - DNA encoding fusion protein consisting of a
portion of HCV polymerase and histidine tag at the
C-terminus

<221> misc_feature
<222> 144, 152, 214, 294, 312, 560, 622, 716, 776, 939, 1007, 1017,
1031, 1095, 1183, 1249, 1270, 1349, 1405, 1408, 1501, 1588,
1604, 1679, 1824, 1899, 2076, 2086, 2221, 2225, 2240, 2313,
2445, 2605, 2634, 2760
<223> n = A,T,C or G

<221> misc_feature
<222> 144, 152, 214, 294, 312, 560, 622, 716, 776, 939, 1007, 1017,
1031, 1095, 1183, 1249, 1270, 1349, 1405, 1408, 1501, 1588,
1604, 1679, 1824, 1899, 2076, 2086, 2221, 2225, 2240, 2313,
2445, 2605, 2634, 2760
<223> n = A,T,C or G

<221> misc_feature
<222> 144, 152, 214, 294, 312, 560, 622, 716, 776, 939, 1007, 1017,
1031, 1095, 1183, 1249, 1270, 1349, 1405, 1408, 1501, 1588,
1604, 1679, 1824, 1899, 2076, 2086, 2221, 2225, 2240, 2313,
2445, 2605, 2634, 2760
<223> n = A,T,C or G

<221> misc_feature
<222> 144, 152, 214, 294, 312, 560, 622, 716, 776, 939, 1007, 1017,
1031, 1095, 1183, 1249, 1270, 1349, 1405, 1408, 1501, 1588,
1604, 1679, 1824, 1899, 2076, 2086, 2221, 2225, 2240, 2313,
2445, 2605, 2634, 2760
<223> n = A,T,C or G

<221> misc_feature
<222> 144, 152, 214, 294, 312, 560, 622, 716, 776, 939, 1007, 1017,
1031, 1095, 1183, 1249, 1270, 1349, 1405, 1408, 1501, 1588,
1604, 1679, 1824, 1899, 2076, 2086, 2221, 2225, 2240, 2313,
2445, 2605, 2634, 2760
<223> n = A,T,C or G

<400> 2
atgtcaatgt cctacacatg gacagggcgc ttgateacgc catgcgetmt armtsrtyrt 60
hrtrthrgya authrrcyaa agcggaggaa agcaagctgc ccatcaacgc gttgagcaac 120
tctttgctga agugusrysu rasnaausra snsrucgcc accataacat ggtttatgcc 180
acaacatctc gcagcgcagg cctgarghsh sasnmtvaty raathrthrs rargxraagy 240
ucggcagang aaggtcacct ttgacagact gcaagtcctg gacgaccaca rggnysysva 300
thrhasargu gnvaussash staccgggac gtgctcaagg agatgaaggc gaaggcgctcc 360
acagttaagt yrargasvau ysgumtyyaa ysaasrthrv aysgctaaac tcctatccgt 420
agaggaagcc tgcaagctga cgcgccccaca taaysuusrv aguguaacys ysauthrrrhs 480

```

tcggccaaat ccaagtttgg ctatggggca aaggacgtcc ggaacctasr aayssryshg 540
ytyrgyaays asvaargasn utccagcaag gccgttaacc acatccactc cgtgtggaag 600
gacttgctgs rarysaavaa snhshsrrva tryasauuga agacactgtg acaccaattg 660
acaccaccat catggcaaaa aatgagguas thrvathrra sthrthrmta aysasnugt 720
ttctctgtgc caaccagaga aaggaggccg taagccagcc cgccttvahc ysvagnrguy 780
sgygyargys raaarguac gtattcccag atctgggagt cgtgtgatgc gagaagatgg 840
ccctcvahra sugyvaargv acysguysmt aautatgatg tggctctccac ccttccctcag 900
gtcgtgatgg gctccrcata ctyrasvava srthrurgnv avamcgysrs rtyrggattc 960
cagtactetc ctgggcagcg agtcgagttc ctggtgaata ccgyhgntyr srrgygnarg 1020
vaguhuvaas nthrtggaaa tcaaagaaaa accccatggg cttttcatat gacactcgtc 1080
gltryssrys ysasnrmtyg hsrtyrasth rargcysttc gactcaacgg tcaccagagaa 1140
cgacatccgt gttgaggagt caatthasr thrvathrgu asnasargva gugusrtacc 1200
aatggttgga ctgggcccc gaagccagac agggccataaa atcgtyrngc yscysasuaa 1260
rguaaarggn aayssrctca cagagcggct ttatatcggg ggtcctctga ctaattcaaa 1320
aggguthrgu argutyrgyg yruthrasns rysgycagaa ctgcggttat cgcgggtgac 1380
gcgcgagcgg cgtgctgacg actgnasncy sgytyrarga rgcysargaa srgyvauthr 1440
thragctgcg gtaacaccct cacatgttac ttgaaggcct ctgcagcctg tsrccsgyas 1500
nthruthrcy atyruysaas raaaacyscg agctgcgaag ctccaggact gcacgatgct 1560
cgtgaacgga gacgacarga aaaysugnas cysthrmtuv aasngyasas ctcgctgcta 1620
tctgtgaaag cgcgggaacc caagaggacg cggcgcagcuv avacysgusr aagythrgng 1680
uasaanaasrc tacgagtctt cacggaggct atgactaggt actccgcccc ccccgguuar 1740
gvahthrgua ambthrargt yraraarrgy gacccgcccc aaccagaata cgacttgag 1800
ctgataacat catgttccas rrnrgutytr asuguuthrs rcysrctcca atgtgtcgt 1860
cgcccaacgat gcacacggca aaagggtgta ctacsransv asrvaaahsa saasrgyya 1920
rgvatyrtyr ctcacccgtg atcccaccac cccctcgca cgggctgctg gggagacaut 1980
hrargaarth rthrruaar gaaaatrgut hrctagaca cactccagtt aactcctggc 2040
taggcaacat tattatgtat aaarghsthr rvaasnrrtr ugyasnmtty rgcgccact 2100
ttgtgggcaa ggatgattct gatgactcac ttcttctcca arthrutraa argmtumtth 2160
rhshhsratc cttctagcgc aggagcaact tgaaaaagcc ctggactgcc agatcuuaag 2220
ngugnuguya nauascysgn tacggggcct gttactccat tgagccactt gacctacctc 2280
agatcattty rgyaacysty rarguruasu rgngaacgac tccatggcct tagcgattt 2340
tcactccata gttactctcc aguarguhsg yusraahsru hsrtyrsrr ggtgagatca 2400
atagggtggc ttcctgcttc aggaacttg ggggtaccag guasnargva aasrcysuar 2460
gybugyvarc ccttgcgagt ctggagacat cgggcccagg gcgtccgcgc taggctaru 2520
rgvatrargh sargaaargv rvaargaaar guctgtccca gggggggagg gccgccactt 2580
gtggcaagta cctcttcaac usrgngygya rgaaanthrc ysgyystyru hasntgggca 2640
gtgaagacca aactcaaac cactccaatc cggctgctg cctraavays thrysuyaut 2700
hrrraaasr cagctggact tgtccggctg gttcgttgc ggttacagcg ggggagacgn 2760
uasusrgytr hvaaagytyr srgygyasat atatcacagc ctgtctcgtg cccgaccccg 2820
cggatcccat caccattyth sersurarga aargrarggy srhshshsca ccatcactaa 2880
cahshshshs

```

<210> 3

<211> 579

<212> PRT

<213> Artificial Sequence

<220>

<223> DNA encoding fusion protein consisting of a
portion of HCV polymerase and histidine tag at the
C-terminus

<400> 3

Met Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala

1 5 10 15
 Ala Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu
 20 25 30
 Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala Gly Leu
 35 40 45
 Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Asp His
 50 55 60
 Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Ala Ser Thr Val Lys
 65 70 75 80
 Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro Pro His
 85 90 95
 Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Asn Leu
 100 105 110
 Ser Ser Lys Ala Val Asn His Ile His Ser Val Trp Lys Asp Leu Leu
 115 120 125
 Glu Asp Thr Val Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu
 130 135 140
 Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu
 145 150 155 160
 Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu
 165 170 175
 Tyr Asp Val Val Ser Thr Leu Pro Gln Val Val Met Gly Ser Ser Tyr
 180 185 190
 Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Thr
 195 200 205
 Trp Lys Ser Lys Lys Asn Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys
 210 215 220
 Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile
 225 230 235 240
 Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Lys Ser
 245 250 255
 Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly
 260 265 270
 Gln Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr
 275 280 285
 Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys
 290 295 300
 Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Asn Gly Asp Asp
 305 310 315 320
 Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser
 325 330 335
 Leu Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly
 340 345 350
 Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser
 355 360 365
 Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr
 370 375 380
 Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr
 385 390 395 400
 Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr
 405 410 415
 Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser
 420 425 430
 Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Gln Ile

435 440 445
 Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile
 450 455 460
 Glu Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro
 465 470 475 480
 Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro
 485 490 495
 Pro Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu
 500 505 510
 Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn
 515 520 525
 Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser
 530 535 540
 Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp
 545 550 555 560
 Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Gly Ser His His His
 565 570 575
 His His His

<210> 4
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer_bind - Artificially synthesized primer
 sequence, 5BNDc1FW

<400> 4
 catatgtcaa tgtcctacac atggacagcc

30

<210> 5
 <211> 57
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer_bind - Artificially synthesized primer
 sequence, 5B570HRV

<400> 5
 ttattagtga tggatgatggt gatgggatcc gcggggtcgg gcacgagaca ggctgtg

57

<210> 6
 <211> 57
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer_bind - Artificially synthesized primer
 sequence, 5B552HRV

<400> 6
ttattagtga tggatgatggt gatgggatcc aacgaaccag ccggacaagt ccagctg 57

<210> 7
<211> 57
<212> DNA
<213> Artificial Sequence

<220>
<223> primer_bind - Artificially synthesized primer
sequence, 5B544HRV

<400> 7
ttattagtga tggatgatggt gatgggatcc ctgggaacga gccgggattg gagtga 57

<210> 8
<211> 67
<212> DNA
<213> Artificial Sequence

<220>
<223> primer_bind - Artificially synthesized primer
sequence, 5B536HRV

<400> 8
ttattagtga tggatgatggt gatgggatcc gagtttgagt ttggtcttca ctgccagtt 60
gaagagg 67

<210> 9
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> primer_bind - Artificially synthesized primer
sequence, 5B531HRV

<400> 9
ttattagtga tggatgatggt gatgggatcc cttcactgcc cagttgaaga ggtacttgcc 60

<210> 10
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<223> primer_bind - Artificially synthesized primer
sequence, 5B591HRV

<400> 10
ttattaatgg tgatggatgat ggtgtccgga tcgattgggg agcaggtaga tg 52

<210> 11

<211> 8
<212> PRT
<213> Hepatitis C virus

<220>
<221> VARIANT
<222> 1, 8
<223> Xaa = Any Amino Acid

<400> 11
Xaa Asp Leu Ser Gly Trp Phe Xaa
1 5

B1
cont

<210> 12
<211> 8
<212> PRT
<213> Hepatitis C virus

<400> 12
Lys Asp Leu Ser Gly Trp Phe Lys
1 5